I picked Mlb salary data, I picked this set because it was the most current one I could find, author unknown. This is relevant because it shows who gets paid the most and who is making the least amount of money. There is no author of this data but is was pulled from the Mlb website database and uploaded to the website I found it on. <https://www.kaggle.com/datasets/kaito510/mlb-player-salaries>

<https://www.spotrac.com/mlb/>

The data was compiled from Kaggle and sporttrac

People would want to look at this because it shows how much people get paid

The columns were describing player salary, names, Team name and player id #. These are all important and need to be separated, for example the player salary Is in its own column because it is in dollar amounts. Player names need to be separated because it is showing who and how much they get paid. The player id is in the last column it is important because if the name was not there or recognizable you could look up their id # and they would pop up. The team name is also important because it is sorted by teams and their name so all the players from the same team will be in sections.

I got new mlb data from <https://github.com/glamp/mlb-payroll-and-wins/blob/master/mlb-standings-and-payroll.csv> this helps because it gives the attendance number’s, average per game and even how long it has been and the estimated payroll. This helps having more variables to be able to make new charts.